DI (DEREK) JIN

Mobile: (919)-813-8126 ⋄ Email: dijin@umich.edu ⋄ URL: `https://derekdijin.github.io/`

Google Scholar: `https://scholar.google.com/citations?user=WwZ7biAAAAAJ`

EDUCATION

University of Michigan, Ann Arbor, MI *2021.11*
PhD in Computational Science & Engineering
Advisor: Danai Koutra

Carnegie Mellon University, Pittsburgh, PA *2016.05*
Master in Computational Data Science (MCDS), School of Computer Science
Research advisor: Christos Faloutsos

Duke University, Durham, NC *2014.05*
Master in Electrical & Computer Engineering

Beijing Institute of Technology (BIT), Beijing *2012.07*
Bachelor of Engineering in Automation

PROFESSIONAL SUMMARY

Applied scientist II, Amazon. Alexa AI-NLU modeling Solutions. *2022.03 - present*

- **Alexa LLM Migration.** Implemented the in-context learning algorithms for the conversational entertainment content discovery for fireTV based on LLM, supporting the LLM-powered Alexa experiences migration.

- **FLARE Expat.** Proposed and implemented a cross-locale training algorithm to reduce the rewrite defects of Alexa utterances transcription. The algorithm learns and transfers the pre-computed rewrites from the US locale with improved user experiences to the international locales. Achieved 26 bps improvement in defect reduction of utterance processing.

- **DERS Rerouting.** Proposed and implemented the Alexa entertainment arbitration model to incorporate customized information and temporal signals. Leveraged the personalized knowledge graph to incorporate the customer temporal interaction and domain information into the distributed classification model. Achieved $\sim$ 2% increase in rerouting performance.

EXPERTISE

- **Research** Large scale graph mining, representation learning, multi-modal fusion, knowledge graphs, deep learning, and social network analysis.

- **Engineering & Analytics** Large-scale data analytics, streaming data analytics, distributed machine learning systems, Pyspark, PySQL, PyTorch, AWS.

PAST WORK EXPERIENCE

- Research intern, Amazon, Seattle. Mentors: Bunyamin Sisman and Luna Xin Dong. *2020.06 - 11*
- Research intern, Tencent, Shenzhen. Mentor: Ying Shan. *2020.03 - 05*
- Research intern, Adobe Research, San Jose. Mentors: Sungchul Kim and Ryan Rossi. *2019.05 - 08*
- Research intern, Adobe Research, San Jose. Mentor: Ryan Rossi. *2018.05 - 08*

SELECTED PUBLICATIONS

- Fatemeh Vahedian, Ruiyu Li, Puja Trivedi, **Di Jin** and Danai Koutra. Leveraging the Graph Structure of Neural Network Training Dynamics. 31st ACM International Conference on Information and Knowledge Management (**CIKM**), 2022.

- **Di Jin**, Bunyamin Sisman, Hao Wei, Luna Xin Dong and Danai Koutra. Deep Transfer Learning for Multi-source Entity Linkage via Domain Adaptation. Proceedings of the VLDB Endowment (**PVLDB**), 2022.

- **Di Jin**, Ryan Rossi, Sungchul Kim and Danai Koutra. On Generalizing Static Node Embedding to Dynamic Settings. The Fifteenth International Conference on Web Search and Data Mining (**WSDM**), Phoenix, AZ, USA, Feb. 2022.

- Nishil Talati* , **Di Jin***, Haojie Ye, Ajay Brahmakshatriya, Saman Amarasinghe, Trevor Mudge, Danai Koutra, and Ronald Dreslinski. A Deep Dive Into Understanding The Random Walk-Based Temporal Graph Learning. The the 2021 IEEE International Symposium on Workload Characterization (**IISWC**), Virtual conference, Nov. 2021.

- **Di Jin**, Bunyamin Sisman, Hao Wei, Luna Xin Dong and Danai Koutra. Deep Transfer Learning for Multi-source Entity Linkage. Amazon Machine Learning Conference (**AMLC**), 2021, (oral paper - 10% acceptance rate)

- **Di Jin**, Yingmin Luo, Zhongang Qi and Ying Shan. TransFusion: Multi-Modal Fusion for Video Tag Inference via Translation-based Knowledge Embedding. ACM MultiMedia (**ACM MM**), Chengdu, China, Oct. 2021.

  Transactions on Knowledge Discovery from Data (**TKDD**) 2020.

- Ryan Rossi, **Di Jin**, Sungchul Kim, Nerseen Ahmed, John Boaz Lee and Danai Koutra. From Community to Role-based Graph Embeddings. Transactions on Knowledge Discovery from Data (**TKDD**) pp. 36. 2020.

- **Di Jin**, Mark Heimann, Ryan Rossi and Danai Koutra. node2bits: Compact Time- and Attribute-aware Node Representations for User Stitching. The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (**ECML-PKDD**), Würzburg, Germany, Sep. 2019.

- **Di Jin**, Ryan Rossi, Eunyee Koh, Sungchul Kim, Anup Rao and Danai Koutra. Latent Network Summarization: Bridging Network Embedding and Summarization. ACM SIGKDD Conference of Knowledge Discovery and Data Mining (**KDD**), Anchorage, Aug. 2019.

- **Di Jin**, Mark Heimann, Tara Safavi, Mengdi Wang, Lee Wei, Lindsay Snider and Danai Koutra. Smart Roles: Inferring Professional Roles in Email Networks. ACM SIGKDD Conference of Knowledge Discovery and Data Mining (**KDD**), Anchorage, Aug. 2019.

- Yujun Yan, Mark Heimann, **Di Jin** and Danai Koutra. Fast Flow-based Random Walk with Restart in a Multi-query Setting. SIAM International Conference on Data Mining (**SDM**), San Diego, May 2018.

- **Di Jin** and Danai Koutra. Exploratory Analysis of Graph Data by Leveraging Domain Knowledge. IEEE International conference of data mining. (**ICDM**), New Orleans, November 2017.

- **Di Jin**, Aristotelis Leventidis, Haoming Shen, Ruowang Zhang, Junyue Wu and Danai Koutra. PERSEUS-HUB: Interactive and Collective Exploration of Large-Scale Graphs. **Informatics** 2017, 4, 22. **(deployed system on Amazon AWS and Azure)**

PATENTS

- Temporal-Based Network Embedding and Prediction   USPTO App. #16/192,313

- Latent Graph Summarization for User Stitching and Online Anomaly Detection   USPTO App. #16/252,169

- A video tag inference method based on multi-modal information fusion   CN 114329062 A

SKILLS

PyTorch, Hadoop, Spark, Java, Python, Matlab, R, Scala, C/C++, Javascript, Bash Script, LaTeX

SERVICE

- PC member and reviewer for **KDD**, **NIPS**, **CIKM**, **WWW**, **WSDM**, **SDM**, **PKDD**

- Reviewer for **Applied Network Science Journal**, **MDPI**, **Neuralcomputing**, open access journal, **PLOS ONE Journal**, **TKDD**, **TKDE**.

- Member for **Frontiers in Big Data** Editorial Board.